stochastic
applications

ELSEVIER Stochastic Processes and their Applications 91 (2001) 239–254
www.elsevier.com/locate/spa

On the small time large deviations of diffusion processes on configuration spaces

T.S. Zhang *

Department of Mathematics, Faculty of Science, HIA, 4604 Kristiansand, Norway
Department of Mathematics, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Received 26 October 1999; received in revised form 31 May 2000; accepted 1 June 2000

Abstract

In this paper, we establish a small time large deviation principle for diffusion processes on configuration spaces. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Dirichlet forms; Intrinsic metric; Large deviations; Configuration spaces

1. Introduction

Let Γ_R be the space of all integer-valued Radon measures on R. The Poisson system of independent Brownian particles can be formally written as

$$X_t = \sum_{i=1}^{\infty} \delta_{B_t^i},$$

where $\{B_t^i,\ i \geqslant 1\}$ are independent Brownian motions. One way to construct the diffusion is to use the Dirichlet form theory (see, for example, Albeverio et al., 1997). The geometric analysis on configuration space Γ_R was carried out in Albeverio et al. (1996a, 1996b, 1997). The intrinsic metric of the associated Dirichlet form was identified as the L^2-Wasserstein-type distance in Röckner and Schied (1999). In this paper, we will establish a small time large deviation principle for the diffusion process $X_t,\ t \geqslant 0$ starting from an explicitly described invariant set. The rate function turns out to be the square of the intrinsic metric of the Dirichlet form.

The paper is organized as follows. Section 2 is the framework. In Section 3, we give an explicit expression of an invariant set of the diffusion. The diffusion starting from every configuration in this set never leaves the set. Section 4 is devoted to the upper bound estimates and identification of the rate function. The lower bound estimates is

* Correspondence address: Department of Mathematics, University of Manchester, Oxford Road, Manchester, M13 9PL, UK.

E-mail address: tzhang@math.uci.edu (T.S. Zhang).

0304-4149/01/$ - see front matter
PII: S0304-4149(00)00062-4

put in Section 5. Section 6 is concerned with a small time large deviation of an average form.

Remark 1.1. The paper is written for the case where the base space is R and the motion is Brownian motion, but the method clearly works for R^d and a large class of diffusions. Now, the natural problem is the large deviation principle on the path space over configuration spaces. This will be studied in a forthcoming paper.

2. Framework

Let Γ_R be the space of all integer-valued Radon measures on the real line R. Then equipped with the vague topology Γ_R is a Polish space. The set of all $\gamma \in \Gamma_R$ such that $\gamma(\{x\}) \in \{0,1\}$ is called configuration space over R. For simplicity, we also call Γ_R configuration space. The geometric analysis on configuration space has been carried out in Albeverio et al. (1997). We recall here some results there. For $f \in C_0(R)$ (the set of all continuous functions on R having compact support), set

$$\langle f, \gamma \rangle = \int_R f(x)\gamma(\mathrm{d}x) = \sum_{x \in \gamma} f(x).$$

Define the space of smooth cylinder functions on Γ_R, $\mathscr{F}C_b^\infty$, as the set of all functions on Γ_R of the form

$$u(\gamma) = F(\langle f_1, \gamma \rangle, \ldots, \langle f_n, \gamma \rangle), \quad \gamma \in \Gamma_R, \tag{2.1}$$

for some $n \in N$, $F \in C_b^\infty(R^n)$, and $f_1, \ldots, f_n \in C_0^\infty(R)$. For u as in (2.1) define its gradient ∇u as a mapping from $\Gamma_R \times R$ to R:

$$\nabla u(\gamma, x) := \sum_{i=1}^n \partial_i F(\langle f_1, \gamma \rangle, \ldots, \langle f_n, \gamma \rangle) \nabla f_i(x), \quad \gamma \in \Gamma_R,\ x \in R.$$

Here ∂_i denotes the partial differential with respect to the ith coordinate, and ∇ is the usual gradient on R.

Let π denote the Poisson measure on Γ_R with intensity $\mathrm{d}x$, i.e., the unique measure on Γ_R whose Laplace transform is given by

$$\int_{\Gamma_R} \mathrm{e}^{\langle f, \gamma \rangle} \pi(\mathrm{d}\gamma) = \exp\left(\int_R (\mathrm{e}^{f(x)} - 1)\mathrm{d}x \right)$$

for all $f \in C_0(R)$. Now we introduce a pre-Dirichlet form:

$$\mathscr{E}_0(u, v) = \int_{\Gamma_R} \langle \nabla u, \nabla v \rangle_\gamma \pi(\mathrm{d}\gamma), \quad u, v \in \mathscr{F}C_b^\infty, \tag{2.2}$$

where $\langle \nabla u, \nabla v \rangle_\gamma = \int_R \nabla u(\gamma, x) \cdot \nabla v(\gamma, x) \gamma(\mathrm{d}x)$. It has been shown in Röckner and Schied (1999) that $\mathscr{E}_0(u, v) = \int_{\Gamma_R} \langle \nabla u, \nabla v \rangle_\gamma \pi(\mathrm{d}\gamma)$ is closable on $L^2(\Gamma_R, \pi)$ and its closure, denoted by $(\mathscr{E}, D(\mathscr{E}))$, is a quasi-regular Dirichlet form. Thus, according to the general theory of Dirichlet forms (see Ma and Röckner (1992)), there is a diffusion process $\{\Omega, X_t, P_\gamma,\ \gamma \in \Gamma_R\}$ associated with the Dirichlet form $(\mathscr{E}, D(\mathscr{E}))$. It was proved in Albeverio et al. (1997) that the diffusion X_t can be identified as infinitely many independent Brownian particles, i.e.,

$$X_t = \sum_i \delta_{B_t^i}, \quad t \geqslant 0,$$

where B_t^i, $i = 1, 2, 3, \dots$ are independent Brownian motions, and δ_x denotes the Dirac measure at the point x. The diffusion X itself is also called Brownian motion on the configuration space.

For $u \in D(\mathscr{E})$, set $\Gamma(u,u)(\gamma) = \langle \nabla u, \nabla u \rangle_\gamma$. Recall that the intrinsic metric of the Dirichlet form $(\mathscr{E}, D(\mathscr{E}))$ is defined by

$$\rho(\gamma,\eta) = \sup\{u(\gamma) - u(\eta); u \in D(\mathscr{E}) \cap C(\Gamma_R) \quad \text{and} \quad \Gamma(u,u) \leqslant 1 \ \pi\text{-a.e. on } \Gamma_R\}$$

for $\gamma, \eta \in \Gamma_R$. It was shown in Röckner and Schied (1999) that the intrinsic metric is given by the Wasserstein-type distance on Γ_R.

$$\rho(\gamma,\eta) = \inf\left\{ \sqrt{\int_{R\times R} d(x,y)^2 \omega(\mathrm{d}x, \mathrm{d}y)}; \omega \in \Gamma_{\gamma\times\eta} \right\}, \tag{2.3}$$

where $\Gamma_{\gamma\times\eta}$ denotes the set of $\omega \in \Gamma_{R\times R}$ having marginals γ and η, and $d^2(x,y) = \frac{1}{2}|x-y|^2$.

3. Identification of the configurations

The Dirichlet form theory tells us that the diffusion $\{\Omega, X_t, P_\gamma, \gamma \in \Gamma_R\}$ exists for quasi-every starting points, but does not give us the exceptional set explicitly. It is easy to see that the diffusion process X_t, $t > 0$ can not start from every configuration $\gamma \in \Gamma_R$, for example, $X_t = \sum_i \delta_{B_t^i}$ with initial value $\gamma = \sum_{i=1}^\infty \delta_{\ln(i)}$ will not be a Radon measure on R. In this section, we will identify the exceptional set and find an explicit invariant set for the diffusion. Let us first recall the definition of capacity. For an open subset G, the capacity is defined as

$$\mathrm{Cap}(G) = \inf\{\mathscr{E}_1(u,u);\ u \in D(\mathscr{E}),\ u \geqslant 1 \ \pi\text{-a.e. on } G\},$$

where $\mathscr{E}_1(u,u) = \mathscr{E}(u,u) + (u,u)_{L^2(\Gamma_R,\pi)}$.

For an arbitrary subset A,

$$\mathrm{Cap}(A) = \inf\{\mathrm{Cap}(G); A \subset G, G \text{ open}\}.$$

Let $g(x) = 1/(1+x^2)$. Define $\Gamma_g = \{\gamma; \langle g, \gamma \rangle < \infty\}$. We have

Lemma 3.1. *$Cap(\Gamma_g^c) = 0$.*

Proof. Set $u(\gamma) = \langle g, \gamma \rangle$. The lemma follows if we show that u is quasi-continuous and belongs to $D(\mathscr{E})$, since a quasi-continuous version of an element in the domain of the Dirichlet form is finite quasi-everywhere. First we note that

$$\int_{\Gamma_R} u(\gamma)^2 \pi(\mathrm{d}\gamma) = \left(\int_R g(x)\,\mathrm{d}x\right)^2 + \int_R g^2(x)\,\mathrm{d}x < \infty.$$

Furthermore,

$$\int_{\Gamma_R} \langle \nabla u, \nabla u \rangle_\gamma \pi(\mathrm{d}\gamma) = \int_{\Gamma_R} \langle g'(x)^2, \gamma \rangle \pi(\mathrm{d}\gamma) = \int_R \frac{4x^2}{(1+x^2)^4}\,\mathrm{d}x < \infty.$$

Combining these two yields $u \in D(\mathscr{E})$. Now, take a sequence of functions $\phi_n \in C_0^\infty(R)$, $n \geqslant 1$, such that $0 \leqslant \phi_n \leqslant 1$, $|\phi_n'(x)| \leqslant 1$, $\phi_n(x) = 1$ on $[-n, n]$ and 0 on $(-\infty, -n-1] \cup [n+1, \infty)$. Set $g_n(x) = g(x)\phi_n(x)$. Define $u_n(\gamma) = \langle g_n, \gamma \rangle$, for $n \geqslant 1$. Then it is easy to see that $u_n(\gamma)$ converges to $u(\gamma)$ in the Dirichlet space and also pointwisely. Since u_n, $n \geqslant 1$, is continous, it follows from Theorem 2.1.4. in Fukushima et al. (1994) that u is quasi-continuous.

Lemma 3.2. *Γ_g is an invariant set for the diffusion $\{\Omega, X_t, P_\gamma, \gamma \in \Gamma_R\}$, i.e.,*

$$P_\gamma(X_t \in \Gamma_g,\ \forall t \geqslant 0) = 1, \quad \forall \gamma \in \Gamma_g.$$

Proof. Fix $\gamma \in \Gamma_R$ with $\langle g, \gamma \rangle = \sum_{i=1}^\infty 1/(1+(\gamma^i)^2) < \infty$. We need to show that

$$P_\gamma\left(\sum_{i=1}^\infty \frac{1}{1+(B_t^i)^2} < \infty,\ \forall t \geqslant 0\right) = 1.$$

By the choice of γ, there exists an integer m_0 such that $\gamma^i \neq 0$ for $i \geqslant m_0$. Since under P_γ, B_t^i, $i \geqslant 1$ are independent Brownian motions with initial values γ^i, $i \geqslant 1$, we have for any integer n,

$$\sum_{i=m_0}^\infty P_\gamma\left(\sup_{s \leqslant n} |B_s^i - \gamma^i| > \frac{|\gamma^i|}{3}\right) \leqslant \sum_{i=m_0}^\infty P_\gamma\left[\sup_{s \leqslant n} |B_s^i - \gamma^i|^2\right] \frac{1}{((1/3)\gamma^i)^2} < \infty,$$

where we used the fact that $P_\gamma[\sup_{s \leqslant n} |B_s^i - \gamma^i|^2]$ is independent of i. By the Borel–Cantelli lemma it follows that

$$P_\gamma\left(\bigcap_k \bigcup_{i \geqslant k} \left(\sup_{s \leqslant n} |B_s^i - \gamma^i| > \frac{|\gamma^i|}{3}\right)\right) = 0.$$

Define

$$\Omega' = \bigcap_n \bigcup_k \bigcap_{i \geqslant k} \left(\sup_{s \leqslant n} |B_s^i - \gamma^i| \leqslant \frac{|\gamma^i|}{3}\right).$$

Then, $P_\gamma(\Omega') = 1$. The lemma follows since

$$\Omega' \subset \left(\sum_{i=1}^\infty \frac{1}{1+(B_t^i)^2} < \infty,\ \forall t \geqslant 0\right).$$

4. Large deviation estimates: the upper bound

Let χ denote the set of all signed Radon measures on R. Equip χ with the vague topology generated by

$$\{U_{f,x} = \{\nu \in \chi; |\langle f, \nu \rangle - x| < \delta\}, f \in C_0(R),\ x \in R,\ \delta > 0\}.$$

Then, χ is a locally convex topological space with χ^* being identified as $C_0(R)$. It is known that Γ_R is a closed subspace of χ. Let $\gamma_0 \in \Gamma_g$ be fixed. Define, for $\eta \in \chi$,

$$I_{\gamma_0}(\eta) = \sup_{f \in C_0(R)} \left(\langle f, \eta \rangle - \int_R \sup_y \left(f(y) - \frac{1}{2}|x-y|^2 \right) \gamma_0(\mathrm{d}x) \right). \tag{4.1}$$

Lemma 4.1. *Let $f \in C_0(R)$. Then*

$$\Lambda(f) =: \lim_{\varepsilon \to 0} \varepsilon \log P_{\gamma_0}[\mathrm{e}^{(1/\varepsilon)\langle f, X_\varepsilon \rangle}] = \int_R \sup_y \left(f(y) - \frac{1}{2}|x-y|^2 \right) \gamma_0(\mathrm{d}x).$$

Proof. By the independence,

$$\begin{aligned}\Lambda(f) &= \lim_{\varepsilon \to 0} \varepsilon \log \prod_{\gamma_0^i \in \gamma_0} P_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}] \\ &= \lim_{\varepsilon \to 0} \sum_i \varepsilon \log P_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}].\end{aligned}$$

By the large deviation principle of Brownian motion (see, for example, Dembo and Zeitouni, 1992), it follows that

$$\lim_{\varepsilon \to 0} \varepsilon \log P_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}] = \sup_y (f(y) - \tfrac{1}{2}|\gamma_0^i - y|^2).$$

Taking the limit inside the series, we get

$$\Lambda(f) = \sum_i \sup_y \left(f(y) - \frac{1}{2}|\gamma_0^i - y|^2 \right) = \int_R \sup_y \left(f(y) - \frac{1}{2}|x-y|^2 \right) \gamma_0(\mathrm{d}x).$$

It now remains to justify taking the limit inside the series. We suppose that the support of f is contained in $[-a, a]$. Without loss of generality, we can also assume $\gamma_0^i > 2a$. With these assumptions, we have

$$\begin{aligned} E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}] &\leqslant E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon)||f||_\infty \chi_{[-a,a]}(B_\varepsilon)}] \\ &= \mathrm{e}^{(1/\varepsilon)||f||_\infty} \int_{-a}^{a} \mathrm{e}^{-(y-\gamma_0^i)^2/2\varepsilon} \frac{1}{\sqrt{2\pi\varepsilon}}\, \mathrm{d}y + \int_{|y|>a} \mathrm{e}^{-(y-\gamma_0^i)^2/2\varepsilon} \frac{1}{\sqrt{2\pi\varepsilon}}\, \mathrm{d}y \\ &\leqslant 1 + \mathrm{e}^{(1/\varepsilon)||f||_\infty} P(B_\varepsilon < a - \gamma_0^i) \\ &\leqslant 1 + \mathrm{e}^{(1/\varepsilon)||f||_\infty} \mathrm{e}^{-\alpha(a-\gamma_0^i)^2/\varepsilon} C, \end{aligned}$$

where $C = E[\mathrm{e}^{\alpha(B_1)^2}]$ for some $\alpha > 0$.

$$\leqslant 1 + \mathrm{e}^{-\alpha'(\gamma_0^i)^2/\varepsilon} C$$

for big enough i and an appropriate choice of α'. By the Schwartz's inequality,

$$\begin{aligned} (E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}])^{-1} &\leqslant E_{\gamma_0^i}[\mathrm{e}^{-(1/\varepsilon) f(B_\varepsilon)}] \\ &\leqslant E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon)||f||_\infty \chi_{[-a,a]}(B_\varepsilon)}] \leqslant 1 + \mathrm{e}^{-\alpha'(\gamma_0^i)^2/\varepsilon} C. \end{aligned}$$

Hence, it follows that for $\varepsilon \leqslant 1$,

$$|\log E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon) f(B_\varepsilon)}]| \leqslant \log(1 + \mathrm{e}^{-\alpha'(\gamma_0^i)^2/\varepsilon} C) \leqslant \mathrm{e}^{-\alpha'(\gamma_0^i)^2/\varepsilon} C \leqslant \mathrm{e}^{-\alpha'(\gamma_0^i)^2} C.$$

Since $\sum_{i=1}^{\infty} 1/(\gamma_0^i)^2 < \infty$, the above estimates show that $\sum_i \varepsilon \log P_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon)f(B_\varepsilon)}]$ converges uniformly with respect to ε, which justifies the limit inside the series.

Proposition 4.2. *Let μ_ε be the law of X_ε on χ under P_{γ_0}. Then $\{\mu_\varepsilon,\ \varepsilon > 0\}$ is exponentially tight, namely, for any $L > 0$, there exists a compact subset K_L such that*

$$\limsup_{\varepsilon\to 0} \varepsilon \log P_{\gamma_0}(X_\varepsilon \in K_L^{\mathrm{c}}) \leqslant -L. \tag{4.2}$$

Proof. Observe that the set of the following type:

$$K = \bigcap_n \{\mu \in \chi;\ |\mu|([-n,n]) \leqslant L_n\} \tag{4.3}$$

is relatively compact. Given $L > 0$, we are going to choose appropriate L_n so that (4.2) holds. Let us fix a positive number α_0 such that $E[\mathrm{e}^{\alpha_0 |B_1|^2}] < \infty$. Let $m_n = \sharp\{i;\ |\gamma_0^i - n| \leqslant \sqrt{1/\alpha_0}\}$. Recall from the proof of Lemma 4.1 that

$$E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon)\chi_{[-n,n]}(B_\varepsilon)}] \leqslant 1 + C\mathrm{e}^{-(\alpha_0(n-\gamma_0^i)^2-1)/\varepsilon}.$$

Hence,

$$\begin{aligned}
&P_{\gamma_0}(X_\varepsilon([-n,n]) > L_n)\\
&\quad\leqslant \mathrm{e}^{-L_n/\varepsilon} E_{\gamma_0}[\mathrm{e}^{(1/\varepsilon)\sum_{i=1}^{\infty}\chi_{[-n,n]}(B_\varepsilon)}]\\
&\quad= \mathrm{e}^{-L_n/\varepsilon}\prod_{i=1}^{\infty} E_{\gamma_0^i}[\mathrm{e}^{(1/\varepsilon)\chi_{[-n,n]}(B_\varepsilon)}]\\
&\quad\leqslant \mathrm{e}^{-L_n/\varepsilon}\prod_{i=1}^{\infty}(1 + C\mathrm{e}^{-(\alpha_0(n-\gamma_0^i)^2-1)/\varepsilon})\\
&\quad\leqslant \mathrm{e}^{-L_n/\varepsilon}\prod_{\alpha_0(\gamma_0^i-n)^2-1>0}(1 + C\mathrm{e}^{-(\alpha_0(n-\gamma_0^i)^2-1)})\prod_{|\gamma_0^i-n|\leqslant\sqrt{1/\alpha_0}}(1 + C\mathrm{e}^{1/\varepsilon})\\
&\quad\leqslant \mathrm{e}^{-L_n/\varepsilon}\exp\left\{C\sum_i \mathrm{e}^{-(\alpha_0(n-\gamma_0^i)^2-1)}\right\}(1 + C\mathrm{e}^{1/\varepsilon})^{m_n}\\
&\quad\leqslant \mathrm{e}^{-L_n/\varepsilon}\exp\left\{C\mathrm{e}^{\alpha_0 n^2+1}\sum_i \mathrm{e}^{-((1/2)\alpha_0(\gamma_0^i)^2)}\right\}(2C)^{m_n}\mathrm{e}^{m_n/\varepsilon}\\
&\quad= \exp\left\{-\frac{L_n - m_n}{\varepsilon} + C_0\mathrm{e}^{\alpha_0 n^2} + m_n\log(2C)\right\},
\end{aligned}$$

where $C_0 = C\mathrm{e}\sum_i \mathrm{e}^{-(1/2)\alpha_0(\gamma_0^i)^2}$. Now take $L_n = L + m_n + C_0\mathrm{e}^{\alpha_0 n^2} + m_n\log(2C) + n$ and define K_L as in (4.3). We have

$$P_{\gamma_0}(X_\varepsilon \notin K_L) \leqslant \sum_{n=1}^{\infty} P_{\gamma_0}(X_\varepsilon([-n,n]) > L_n)$$

$$\leqslant \sum_{n=1}^{\infty} c \exp\left\{-\frac{L_n - m_n}{\varepsilon} + C_0 e^{\alpha_0 n^2} + m_n \log(2C)\right\}$$

$$\leqslant \sum_{n=1}^{\infty} \exp\left\{-\frac{L+n}{\varepsilon}\right\} \leqslant C e^{-L/\varepsilon}.$$

This implies (4.2), which ends the proof. □

Combination of Lemma 4.1, Proposition 4.2 and Theorem 4.5.20 in Dembo and Zeitouni (1992) yields the following upper bound estimates:

Theorem 4.3. *Let μ_ε be the law of X_ε under P_{γ_0}. Then, for any closed subset $F \subset \Gamma_R$,*

$$\limsup_{\varepsilon \to 0} \varepsilon \log \mu_\varepsilon(F) \leqslant -\inf_{\eta \in F} I_{\gamma_0}(\eta). \tag{4.4}$$

Next, we are going to identify the rate function $I_{\gamma_0}(\eta)$ as $\rho(\gamma_0, \eta)^2$. Let $M(R)$ denote the space of positive Radon measures on R. Set $d(x,y) = (1/\sqrt{2})|x-y|$. Let $\omega, \gamma \in M(R)$. The L^2-Wasserstein-type distance $\rho(\omega, \gamma)$ is similarly defined as

$$\rho(\omega,\gamma) = \inf\left\{\sqrt{\int_{R\times R} d(x,y)^2 \eta(dx,dy)};\ \eta \in \Gamma_{\omega,\gamma}\right\},$$

where $\Gamma_{\omega,\gamma}$ stands for the set of $\eta \in M(R\times R)$ having marginals ω and γ.

Now fix $\gamma_0 \in M(R)$. Define $h(\gamma) = \rho(\gamma_0, \gamma)^2$.

Proposition 4.4. *$h(\gamma)$ is convex and lower semi-continuous on $M(R)$ with the topology of vague convergence.*

Proof. Let $\gamma_1, \gamma_2 \in M(R)$. For any $0 < \alpha < 1$, $\eta^1 \in \Gamma_{\gamma_0,\gamma_1}$, $\eta^2 \in \Gamma_{\gamma_0,\gamma_2}$, define $\eta = \alpha\eta^1 + (1-\alpha)\eta^2 \in M(R\times R)$. Then,

$$\eta(dx,R) = \alpha\eta^1(dx,R) + (1-\alpha)\eta^2(dx,R) = \alpha\gamma_0(dx) + (1-\alpha)\gamma_0(dx) = \gamma_0(dx),$$

$$\eta(R,dy) = \alpha\gamma_1(dy) + (1-\alpha)\gamma_2(dy).$$

Hence $\eta \in \Gamma_{\gamma_0,\, \alpha\gamma_1 + (1-\alpha)\gamma_2}$, and

$$\begin{aligned} h(\alpha\gamma_1 + (1-\alpha)\gamma_2) &\leqslant \int_{R\times R} d^2(x,y)\eta(dx,dy) \\ &= \alpha \int_{R\times R} d^2(x,y)\eta^1(dx,dy) + (1-\alpha)\int_{R\times R} d^2(x,y)\eta^2(dx,dy). \end{aligned} \tag{4.5}$$

Since η^1 and η^2 are arbitrary, it follows from (4.5) that

$$h(\alpha\gamma_1 + (1-\alpha)\gamma_2) \leqslant \alpha h(\gamma_1) + (1-\alpha)h(\gamma_2)$$

which completes the proof of the convexity. Next, we prove that h is lower semi-continuous. Let $\{\gamma_n,\ n \geqslant 1\}$ be a sequence of positive Radon measures on R converging vaguely to $\bar{\gamma} \in M(R)$. We need to show

$$h(\bar{\gamma}) \leqslant \liminf_{n\to\infty} h(\gamma_n). \tag{4.6}$$

To this end, without loss of generality, we assume that the limit $\lim_{n\to\infty} h(\gamma_n)$ exists and is finite. By the definition of ρ, there exists a sequence of Radon measures $\eta_n(\mathrm{d}x,\mathrm{d}y)\in M(R\times R)$ such that $\eta_n\in\Gamma_{\gamma_0,\gamma_n}$ and

$$\lim_{n\to\infty} h(\gamma_n)=\lim_{n\to\infty}\int_{R\times R} d^2(x,y)\eta_n(\mathrm{d}x,\mathrm{d}y).$$

Let Π_1 be the projection operator from $R\times R$ to R defined by $\Pi_1(x,y)=x$. For any compact set $K\subset R\times R$, we have

$$\sup_n \eta_n(K)\leqslant \sup_n \Pi_1^*\eta_n(\Pi_1(K))=\gamma_0(\Pi_1(K))<\infty.$$

This shows that the family $\{\eta_n,\ n\geqslant 1\}$ is relatively compact with respect to the topology of vague convergence. Let $\{\eta_{n_k},\ k\geqslant 1\}$ be a subsequence converging vaguely to some Radon measure η_0 on $R\times R$. Next, we prove that

$$\eta_0(\mathrm{d}x,R)=\gamma_0(\mathrm{d}x),\quad \eta_0(R,\mathrm{d}y)=\bar{\gamma}(\mathrm{d}y).$$

Note that this is not the consequence of vague convergence. Because of the similarity, we prove one of them, say, $\eta_0(R,\mathrm{d}y)=\bar{\gamma}(\mathrm{d}y)$.

Choose a sequence $\{\phi_m(x),\ m\geqslant 1\}$ of C^∞ functions satisfying $0\leqslant\phi_m(x)\leqslant 1$, $\phi_m(x)=1$ on $[-m,m]$, $\phi_m(x)=0$ on $[-m-1,m+1]^c$. Let any $f\in C_0(R)$. Suppose $\mathrm{Supp}[f]\subset[-m_0,m_0]$ for some m_0. Then,

$$\begin{aligned}\int_{R\times R} f(y)\eta_0(\mathrm{d}x,\mathrm{d}y)&=\lim_{m\to\infty}\int_{R\times R} f(y)\phi_m(x)\eta_0(\mathrm{d}x,\mathrm{d}y)\\&=\lim_{m\to\infty}\lim_{k\to\infty}\int_{R\times R} f(y)\phi_m(x)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y).\end{aligned}\tag{4.7}$$

But,

$$\begin{aligned}&\left|\lim_{k\to\infty}\int_{R\times R} f(y)\phi_m(x)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)-\lim_{k\to\infty}\int_{R\times R} f(y)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\right|\\&\quad=\left|\lim_{k\to\infty}\int_{R\times R} f(y)(\phi_m(x)-1)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\right|\\&\quad\leqslant c\lim_{k\to\infty}\int_{R\times R}\chi_{[-m_0,\,m_0]}(y)\chi_{[-m,m]^c}(x)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\\&\quad\leqslant c\sup_k\int_{\{d^2(x,y)\geqslant(m-m_0)^2\}}\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\\&\quad\leqslant c\frac{1}{(m-m_0)^2}\sup_k\int_{R\times R} d^2(x,y)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\leqslant\frac{M}{(m-m_0)^2}.\end{aligned}\tag{4.8}$$

Thus, it follows from (4.7) and (4.8) that

$$\begin{aligned}\int_{R\times R} f(y)\eta_0(\mathrm{d}x,\mathrm{d}y)&=\lim_{m\to\infty}\lim_{k\to\infty}\int_{R\times R} f(y)\phi_m(x)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)\\&=\lim_{k\to\infty}\int_{R\times R} f(y)\eta_{n_k}(\mathrm{d}x,\mathrm{d}y)=\lim_{k\to\infty}\int_R f(y)\gamma_{n_k}(\mathrm{d}y)=\int_R f(y)\bar{\gamma}(\mathrm{d}y).\end{aligned}$$

Since f is arbitrary, we conclude that $\eta_0(R,\mathrm{d}y)=\bar{\gamma}(\mathrm{d}y)$. Hence,

$$h(\bar{\gamma}) \leqslant \int_{R\times R} d^2(x,y)\eta_0(\mathrm{d}x,\mathrm{d}y)$$

$$\leqslant \lim_{n\to\infty} \int_{R\times R} d^2(x,y)\eta_n(\mathrm{d}x,\mathrm{d}y) = \lim_{n\to\infty} h(\gamma_n).$$

The proof is complete. □

Lemma 4.5. *If $f\in C_0(R)$, then $g(x)=\sup_y(f(y)-\frac{1}{2}|x-y|^2)\in C_0(R)$.*

Proof. Suppose $\mathrm{Supp}[f]\subset[-M,M]$ and $C=\sup_y|f(y)|$. Then, for $|x|>M$,

$$0\leqslant g(x) = \sup_{y\in[-M,M]}(f(y)-\tfrac{1}{2}|x-y|^2)\vee \sup_{|y|>M}(-\tfrac{1}{2}|x-y|^2)$$

$$\leqslant \sup_{|y|\leqslant M}(C-\tfrac{1}{2}|x-y|^2)\vee 0.$$

The lemma follows as $\sup_{|y|\leqslant M}(C-\frac{1}{2}|x-y|^2)<0$ for big enough $|x|$.

Lemma 4.6. *Let $f\in C_0(R)$. Then*

$$\sup_{\gamma\in\Gamma_R}(<f,\gamma>-\rho(\gamma_0,\gamma)^2) = \int_R \sup_{y\in R}\left(f(y)-\frac{1}{2}|x-y|^2\right)\gamma_0(\mathrm{d}x).$$

Proof. Suppose that $\gamma_0=\sum_{i=1}^{\infty}\delta_{\gamma_0^i}$. Due to the compact support, there exists integer N_0 such that $f(\gamma_0^i)=0$ and $\sup_{y\in R}(f(y)-\frac{1}{2}|\gamma_0^i-y|^2)=0$ for $i>N_0$. Let $\gamma\in\Gamma_R$ with $\rho(\gamma_0,\gamma)<\infty$. According to Röckner and Schied (1999), the distance $\rho(\gamma_0,\gamma)$ is reached by some $\eta\in M(R\times R)$ having marginals γ_0 and γ. Let us say $\eta=\sum_{i=1}^{\infty}\delta_{\gamma_0^i,\gamma^i}$. Thus, $\gamma=\sum_{i=1}^{\infty}\delta_{\gamma^i}$ and $\rho(\gamma_0,\gamma)^2=\frac{1}{2}\sum_{i=1}^{\infty}|\gamma_0^i-\gamma^i|^2$. Hence,

$$\langle f,\gamma\rangle-\rho(\gamma_0,\gamma)^2 = \sum_{i=1}^{\infty} f(\gamma^i)-\frac{1}{2}\sum_{i=1}^{\infty}|\gamma_0^i-\gamma^i|^2$$

$$=\sum_{i=1}^{\infty}\left(f(\gamma^i)-\frac{1}{2}|\gamma_0^i-\gamma^i|^2\right)\leqslant \sum_{i=1}^{\infty}\sup_y\left(f(y)-\frac{1}{2}|\gamma_0^i-y|^2\right)$$

$$=\int_R\sup_y\left(f(y)-\frac{1}{2}|x-y|^2\right)\gamma_0(\mathrm{d}x).$$

On the other hand, for any $\varepsilon>0$, there exist y_i, $i=1,2,\ldots,N_0$ such that

$$\left|\sum_{i=1}^{N_0}\left(f(y_i)-\frac{1}{2}|\gamma_0^i-y_i|^2\right)-\int_R\sup_y\left(f(y)-\frac{1}{2}|x-y|^2\right)\gamma_0(\mathrm{d}x)\right|$$

$$=\left|\sum_{i=1}^{N_0}\left(f(y_i)-\frac{1}{2}|\gamma_0^i-y_i|^2\right)-\sum_{i=1}^{N_0}\sup_y\left(f(y)-\frac{1}{2}|\gamma_0^i-y|^2\right)\right|\leqslant\varepsilon.$$

Define $\gamma^i = y_i$ if $i \leqslant N_0$ and $\gamma^i = \gamma_0^i$ if $i > N_0$. Then,

$$\begin{aligned}&\langle f, \gamma \rangle - \rho(\gamma_0, \gamma)^2 \\ &\quad \geqslant \sum_{i=1}^{N_0} f(y_i) - \sum_{i=1}^{N_0} \frac{1}{2}|y_i - \gamma_0^i|^2 \geqslant \int_R \sup_y \left(f(y) - \frac{1}{2}|x-y|^2 \right) \gamma_0(\mathrm{d}x) - \varepsilon.\end{aligned}$$

Hence,

$$\sup_{\gamma \in \Gamma_R} (\langle f, \gamma \rangle - \rho(\gamma_0, \gamma)^2) \geqslant \int_R \sup_y \left(f(y) - \frac{1}{2}|x-y|^2 \right) \gamma_0(\mathrm{d}x) - \varepsilon.$$

Since ε is arbitrary, the proof is complete. □

Proposition 4.7. *$I_{\gamma_0}(\gamma) = \rho(\gamma_0, \gamma)^2$ for all $\gamma \in \Gamma_R$.*

Proof. By Lemma 4.6,

$$I_{\gamma_0}(\gamma) = \sup_{f \in C_0(R)} \left(\langle f, \gamma \rangle - \sup_{\hat{\gamma} \in \Gamma_R} (\langle f, \hat{\gamma} \rangle - \rho(\gamma_0, \hat{\gamma})^2) \right).$$

Since for any $f \in C_0(R)$, $\langle f, \gamma \rangle - \sup_{\hat{\gamma} \in \Gamma_R} (\langle f, \hat{\gamma} \rangle - \rho(\gamma_0, \hat{\gamma})^2) \leqslant \rho(\gamma_0, \gamma)^2$, we have $I_{\gamma_0}(\gamma) \leqslant \rho(\gamma_0, \gamma)^2$. On the other hand, by the general theorem on the inverse legendre transform of convex functions (see, e.g., Gross, 1980) and Proposition 4.4, we obtain

$$\rho(\gamma_0, \gamma)^2 = \sup_{f \in C_0(R)} \left(\langle f, \gamma \rangle - \sup_{\hat{\gamma} \in M(R)} (\langle f, \hat{\gamma} \rangle - \rho(\gamma_0, \hat{\gamma})^2) \right). \tag{4.9}$$

Since $I_{\gamma_0}(\gamma)$ is greater than the right-hand side of (4.9), the proposition follows.

5. Large deviation estimates: the lower bound

Let U be an open neighborhood described by

$$U = \left\{ \gamma \in \Gamma_R; \gamma(\partial W_r) = 0 \ \gamma|_{W_r} = \sum_{i=1}^n \delta_{x_i} \text{ with } \sum_{i=1}^n d(x_i, y_i)^2 < \delta \right\},$$

where $W_r = \{|x| < r\}$.

Proposition 5.1. *Let $\gamma_0 \in \Gamma_g$. Then, for any $\delta_1 > 0$ and distinct integers $i_1, \ldots, i_n$,*

$$\begin{aligned}\liminf_{\varepsilon \to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in U) &\geqslant -\frac{1}{2} \sum_{j=1}^n (y_j - \gamma_0^{i_j})^2 (1 + \delta_1) - \left(\frac{1}{\delta_1} + 1 \right) \delta \\ &\quad - \sum_{k \notin \{i_1, i_2, \ldots, i_n\}} d(\gamma_0^k, W_r^c)^2.\end{aligned}$$

Proof. For any distinct integers i_1, i_2, $i_3,\dots,i_n$, define

$$A_{i_1,i_2,\dots,i_n} = \left\{ B_\varepsilon^{i_1} \in W_r, \dots, B_\varepsilon^{i_n} \in W_r, \sum_{j=1}^n d(B_\varepsilon^{i_j}, y_j)^2 < \delta, \right.$$

$$\left. B_\varepsilon^k \in \bar{W}_r^c,\ k \notin \{i_1,\dots,i_n\} \right\}.$$

Then,

$$\{X_\varepsilon \in U\} = \bigcup_{i_1,i_2,\dots,i_n} A_{i_1,i_2,\dots,i_n}$$

$$P_{\gamma_0}(X_\varepsilon \in U) = \sum_{i_1,i_2,\dots,i_n}^{\infty} P_{\gamma_0}(A_{i_1,i_2,\dots,i_n}).$$

Now,

$$\begin{aligned}
P_{\gamma_0}(A_{i_1,i_2\dots,i_n}) &= P_{\gamma_0}\left(B_\varepsilon^{i_1} \in W_r, \dots, B_\varepsilon^{i_n} \in W_r, \sum_{j=1}^n d(B_\varepsilon^{i_j}, y_j)^2 < \delta \right) \\
&\quad \times \prod_{k \notin \{i_1,i_2,\dots,i_n\}} P_{\gamma_0}(B_\varepsilon^k \in \bar{W}_r^c) \\
&= \int \cdots \int_{\sum_{j=1}^n d^2(x_j,y_j)<\delta,\ x_j \in W_r} (2\pi\varepsilon)^{-n/2} \\
&\quad \times \exp\left(\frac{-\sum_{j=1}^n (x_j - \gamma_0^{i_j})^2}{2\varepsilon} \right) \mathrm{d}x_1\,\mathrm{d}x_2 \cdots \mathrm{d}x_n \\
&\quad \times \prod_{k \notin \{i_1,i_2,\dots,i_n\}} \int_{|x|>r} \frac{1}{\sqrt{2\pi\varepsilon}} \exp\left(-\frac{(x-\gamma_0^k)^2}{2\varepsilon} \right) \mathrm{d}x \\
&= a_\varepsilon^{i_1,i_2,\dots,i_n} \times b_\varepsilon^{i_1,i_2,\dots,i_n}.
\end{aligned}$$

For any $\delta_1 > 0$, we have

$$\sum_{j=1}^n (x_j - \gamma_0^{i_j})^2 \leqslant \left(1 + \frac{1}{\delta_1}\right) \sum_{j=1}^n (x_j - y_j)^2 + (1+\delta_1) \sum_{j=1}^n (y_j - \gamma_0^{i_j})^2.$$

Thus,

$$\begin{aligned}
a_\varepsilon^{i_1,i_2,\dots,i_n} &\geqslant \int \cdots \int_{\sum_{j=1}^n d^2(x_j,y_j)<\delta,\ x_j\in W_r} (2\pi\varepsilon)^{-n/2} \exp\left(-\frac{(1+1/\delta_1)\delta}{2\varepsilon} \right), \\
&\quad \times \exp\left(-(1+\delta_1) \frac{\sum_{j=1}^n (y_j - \gamma_0^{i_j})^2}{2\varepsilon} \right) \mathrm{d}x_1\,\mathrm{d}x_2 \cdots \mathrm{d}x_n.
\end{aligned}$$

Hence,

$$\liminf_{\varepsilon\to 0} \varepsilon \log a_\varepsilon^{i_1, i_2, \dots, i_n} \geqslant -\frac{1}{2}\left(\frac{1}{\delta_1}+1\right)\delta - (1+\delta_1)\frac{1}{2}\sum_{j=1}^{n}(y_j - \gamma_0^{i_j})^2.$$

To treat $b_\varepsilon^{i_1, i_2, \dots, i_n}$, we need the following

Lemma 5.2. *There exists a constant $K > 0$ such that if $|x| > 2r \vee K$, then*

$$P_x(|B_\varepsilon| > r) \geqslant \mathrm{e}^{-C/\varepsilon x^2}, \tag{5.1}$$

where C is a positive constant, B is a Brownian motion starting from x under P_x.

Proof. We first assume $x > 0$. We have

$$\begin{aligned} P_x(|B_\varepsilon| > r) &= P_0(|B_\varepsilon + x| > r) \\ &\geqslant P_0(B_\varepsilon > r - x) \geqslant P_0\left(B_\varepsilon > -\frac{x}{2}\right) = P_0\left(B_1 > -\frac{x}{2\sqrt{\varepsilon}}\right) \\ &= P_0\left(B_1 < \frac{x}{2\sqrt{\varepsilon}}\right) = \int_{-\infty}^{x/2\sqrt{\varepsilon}} \mathrm{e}^{-y^2/2}\frac{1}{\sqrt{2\pi}}\,\mathrm{d}y. \end{aligned}$$

Thus,

$$P_x(|B_\varepsilon| > r)\mathrm{e}^{C/\varepsilon x^2} \geqslant \mathrm{e}^{C/\varepsilon x^2}\int_{-\infty}^{x/2\sqrt{\varepsilon}} \mathrm{e}^{-y^2/2}\frac{1}{\sqrt{2\pi}}\,\mathrm{d}y.$$

It is therefore, sufficient to show

$$F_{C,\varepsilon}(x) = \mathrm{e}^{C/\varepsilon x^2}\int_{-\infty}^{x/2\sqrt{\varepsilon}} \mathrm{e}^{-y^2/2}\frac{1}{\sqrt{2\pi}}\,\mathrm{d}y \geqslant 1 \tag{5.2}$$

for big enough x and sufficiently small ε.

Since $F_{C,\varepsilon}(x)$ is decreasing in ε, we can let $\varepsilon = \frac{1}{4}$. Put

$$F(x) = F_{C,\frac{1}{4}}(x) = \mathrm{e}^{4C/x^2}\int_{-\infty}^{x/2} \mathrm{e}^{-y^2/2}\frac{1}{\sqrt{2\pi}}\,\mathrm{d}y.$$

Since $\lim_{x\to\infty} F(x) = 1$, to show (5.2), it is sufficient to verify $F'(x) \leqslant 0$ for big enough x. Differentiating F, we get

$$F'(x) = e^{4C/x^2}\left[\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2} - \frac{8C}{x^3}\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\,\mathrm{d}y\right],$$

which is clearly negative when x is big enough.

Now, we assume $x < 0$. In this case,

$$\begin{aligned} P_x(|B_\varepsilon| > r) &\geqslant P_0(B_\varepsilon < -r - x) \geqslant P_0\left(B_\varepsilon < -\frac{x}{2}\right) \\ &= P_0\left(B_1 < -\frac{x}{2\sqrt{\varepsilon}}\right). \end{aligned}$$

By the same reason as above, we can choose $\varepsilon = \frac{1}{4}$. Then, for $\varepsilon \leqslant \frac{1}{4}$,

$$P_x(|B_\varepsilon| > r) \geqslant \int_{-\infty}^{-x}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-y^2/2}\,\mathrm{d}y.$$

Define

$$\hat{F}(x) = \mathrm{e}^{4C/x^2} \int_{-\infty}^{-x} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-y^2/2}\, \mathrm{d}y.$$

We see that $\hat{F}(x) = F(-x)$, which implies that $\hat{F}(x) \geqslant 1$ for big enough $|x|$. The proof is complete. □

Lemma 5.3. *We have*

$$\lim_{\varepsilon\to 0} \varepsilon \log b_\varepsilon^{i_1,i_2,\ldots,i_n} = -\sum_{k\notin\{i_1,\ldots,i_n\}} d(\gamma_0^k, W_r^c)^2. \tag{5.3}$$

Proof. From the definition,

$$\varepsilon \log b_\varepsilon^{i_1,i_2,\ldots,i_n} = \sum_{k\notin\{i_1,\ldots,i_n\}} \varepsilon \log P_{\gamma_0^k}(|B_\varepsilon| > r).$$

By the large deviation principle of Brownian motion, we know that

$$\lim_{\varepsilon\to 0} \varepsilon \log P_{\gamma_0^k}(|B_\varepsilon| > r) = -d(\gamma_0^k, W_r^c)^2$$

On the other hand, it follows from Lemma 5.2 that

$$|\varepsilon \log P_{\gamma_0^k}(|B_\varepsilon| > r)| \leqslant \frac{M}{(\gamma_0^k)^2}.$$

Eq. (5.3) follows from the dominated convergence theorem.

Combination of Lemmas 5.3 and 5.2 gives Proposition 5.1.

Theorem 5.4. *Let $\gamma_0 \in \Gamma_g$. Then for any open set $O \subset \Gamma_R$, we have*

$$\liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in O) \geqslant -\inf_{\gamma\in O} \rho(\gamma_0,\gamma)^2.$$

Proof. It is sufficient to prove

$$\liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in U) \geqslant -\rho(\gamma_0,\hat{\gamma})^2$$

for any $\hat{\gamma} \in O$. Fix such $\hat{\gamma} \in O$. Without loss of generality, we can assume that

$$\rho(\gamma_0,\hat{\gamma})^2 = \frac{1}{2}\sum_{i=1}^{\infty} |\hat{\gamma}^i - \gamma_0^i|^2 < \infty.$$

Choose an increasing sequence W_{r_n} of open balls such that $\bigcup W_{r_n} = R$ and

$$\hat{\gamma}|_{W_{r_n}} = \sum_{i=1}^{m_n} \delta_{\hat{\gamma}^i}, \quad \hat{\gamma}(\partial W_{r_n}) = 0.$$

Let δ_m be a sequence of positive numbers converging to zero. Set

$$U_{n,m} = \left\{ \gamma \in \Gamma_R; \gamma(\partial W_{r_n}) = 0 \ \ \gamma|_{W_{r_n}} = \sum_{i=1}^{m_n} \delta_{x_i} \text{ with } \sum_{i=1}^{m_n} d(x_i,\hat{\gamma}^i)^2 < \delta_m \right\}.$$

Since O is open, there exist n_0, m_0 such that $U_{n,m} \subset O$ for $n \geqslant n_0$, $m \geqslant m_0$.

By Proposition 5.1 it holds that for $m \geqslant m_0, n \geqslant n_0$ and any $\delta_1 > 0$,

$$\liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in O) \geqslant \liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in U_{n,m})$$

$$\geqslant -\frac{1}{2}\sum_{i=1}^{m_n}(\hat{\gamma}^i - \gamma_0^i)^2(1+\delta_1) - \frac{1}{2}\left(\frac{1}{\delta_1}+1\right)\delta_m - \sum_{k=m_n+1}^{\infty} d(\gamma_0^k, \bar{W}^c_{r_n})^2.$$

First letting $m \to \infty$ and then $\delta_1 \to 0$, we obtain that

$$\liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in O) \geqslant -\frac{1}{2}\sum_{i=1}^{m_n}(\hat{\gamma}^i - \gamma_0^i)^2 - \sum_{k=m_n+1}^{\infty} d(\gamma_0^k, \bar{W}^c_{r_n})^2. \tag{5.4}$$

By the choice of W_{r_n}, $d(\gamma_0^k, \bar{W}^c_{r_n}) \leqslant d(\gamma_0^k, \hat{\gamma}^k)$ for $k \geqslant m_n + 1$. Hence, it follows from (5.4) that

$$\liminf_{\varepsilon\to 0} \varepsilon \ln P_{\gamma_0}(X_\varepsilon \in O) \geqslant -\frac{1}{2}\sum_{i=1}^{\infty}(\hat{\gamma}^i - \gamma_0^i)^2 = -\rho(\gamma_0, \hat{\gamma})^2.$$

The proof is complete. □

6. Large deviations of an average form

In this section, we will prove a large deviation principle of an average form for the Brownian motion on configuration space. See the related results for other type of diffusions in Aida and Zhang (1999) and Zhang (2000).

Set

$$P_\pi(\cdot) = \int_{\Gamma_R} P_\gamma(\cdot)\,\mathrm{d}\pi.$$

Let B and C be any Borel subsets of Γ_R. Define

$$\rho(B, C) = \max\left\{\operatorname{ess}\inf_{\gamma\in B} \rho(\gamma, C), \operatorname{ess}\inf_{\gamma\in C} \rho(\gamma, B)\right\}.$$

Theorem 6.1. *Let B, C be two Borel subsets with $\pi(B) > 0$, $\pi(C) > 0$. Then*

$$\limsup_{\varepsilon\to 0} \varepsilon \log P_\pi(X_0 \in B,\ X_\varepsilon \in C) \leqslant -\rho(B, C)^2.$$

Proof. We proceed as in Fang (1994) and Fang and Zhang (1999). Without loss of generality, we can assume $\rho(B, C) > 0$. Let λ be any positive number such that $\lambda < \rho(B, C)$. Then, $\operatorname{ess\,inf}_{\gamma\in B}\ \rho(\gamma, C) > \lambda$ or $\operatorname{ess\,inf}_{\gamma\in C}\rho(\gamma, B) > \lambda$. We assume, for example, the later holds. This implies that there exists a Borel set $K \subset C$ with $\pi(K) = \pi(C)$ such that

$$\rho(\gamma, B) > \lambda \quad \text{for all } \gamma \in K. \tag{6.1}$$

Now fix an integer $n > \lambda$. Define $u(\gamma) = \rho(\gamma, B) \wedge n$. Applying Proposition 3.1 in Röckner and Schied (1999), we have $u \in D(\mathscr{E})$. Then, by the Lyons–Zheng's decomposition (see Lyons and Zhang (1996); Lyons and Zheng (1988)), under P_π the following holds.

$$u(X_s) - u(X_0) = \frac{1}{2}M_s^u - \frac{1}{2}(M_t^u(\gamma_t(\omega)) - M_{t-s}^u(\gamma_t(\omega))) \quad \text{for } 0 \leqslant s \leqslant t, \tag{6.2}$$

where M^u is a $\mathscr{F}_t = \sigma(X_s,\ s \leqslant t)$-square integrable martingale with

$$\langle M^u \rangle_t = \int_0^t \Gamma(u,u)(X_s)\,\mathrm{d}s \tag{6.3}$$

and γ_t is the reverse operator such that $X_s(\gamma_t(\omega)) = \omega(t-s),\ 0 \leqslant s \leqslant t$.

Remarking that π is an invariant measure of the diffusion process, it follows that

$$\begin{aligned} P_\pi(X_0 \in B,\ X_\varepsilon \in C) &= P_\pi(X_0 \in B,\ X_\varepsilon \in K) \\ &\leqslant P_\pi(u(X_\varepsilon) > \lambda,\ u(X_0) = 0) \leqslant P_\pi(u(X_\varepsilon) - u(X_0) > \lambda) \\ &= P_\pi(\tfrac{1}{2}(M_\varepsilon^u - M_\varepsilon^u(\gamma_\varepsilon(\omega))) > \lambda) \\ &\leqslant P_\pi(M_\varepsilon^u > \lambda) + P_\pi(-M_\varepsilon^u > \lambda) \\ &\leqslant 4 \int_{\lambda/\sqrt{\varepsilon}}^\infty \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-s^2/2\,\mathrm{d}s}, \end{aligned} \tag{6.4}$$

where we have used the reversibility of the diffusion and $\langle M^u \rangle_t \leqslant t$ which follows from (6.3) and $\Gamma(u,u)(\gamma) \leqslant 1$. Thus we get from (6.4) that

$$\limsup_{\varepsilon \to 0}\ \varepsilon \log P_\pi(X_0 \in B, X_\varepsilon \in C) \leqslant -\lambda^2. \tag{6.5}$$

Letting $\lambda \to \rho(B,C)$, the assertion follows.

Theorem 6.2. *Let B, C be two Borel subsets with $\pi(B) > 0$, $\pi(C) > 0$. If B or C is open, then*

$$\liminf_{\varepsilon \to 0}\ \varepsilon \log P_\pi(X_0 \in B,\ X_\varepsilon \in C) \geqslant -\rho(B,C)^2. \tag{6.6}$$

Proof. Assume, for example, C is open. Let $\lambda = \rho(B,C)$ and $\delta > 0$. Since $\operatorname{ess\,inf}_{\gamma \in B} \rho(\gamma, C) < \lambda + \delta$, there exists a compact subset $K \subset B \cap \Gamma_g$ with $\pi(K) > 0$ such that $\rho(\gamma, C) < \lambda + \delta$ on K. For any $\gamma \in K$, appplying the large deviation estimates in Theorem 5.4 it follows that

$$\liminf_{\varepsilon \to 0}\ \varepsilon \log P_\gamma(X_\varepsilon \in C) \geqslant -\rho(\gamma, C)^2 > -(\lambda+\delta)^2. \tag{6.7}$$

Fix an arbitrary sequence ε_n of positive numbers converging to zero. It suffices to show (6.6) for such sequences. First, we see that (6.7) implies

$$K \subset \bigcup_m \bigcap_{n \geqslant m} \{\gamma; \varepsilon_n \log P_\gamma(X_{\varepsilon_n} \in C) > -(\lambda+\delta)^2\}. \tag{6.8}$$

Thus, there exists m_0 such that $K_0 = \bigcap_{n \geqslant m_0} \{\gamma; \varepsilon_n \log P_\gamma(X_{\varepsilon_n} \in C) > -(\lambda+\delta)^2\} \cap K$ has positive measure. Hence, by Jensen's inequality,

$$\begin{aligned} \liminf_{n \to \infty} \varepsilon_n \log P_\pi(X_0 \in B, X_{\varepsilon_n} \in C) &\geqslant \liminf_{n \to \infty} \varepsilon_n \log P_\pi(X_0 \in K_0, X_{\varepsilon_n} \in C) \\ &= \liminf_{n \to \infty} \varepsilon_n \log \left(\int_{K_0} P_\gamma(X_{\varepsilon_n} \in C)\pi(\mathrm{d}\gamma) \right) \end{aligned}$$

$$\geqslant \liminf_{n\to\infty} \varepsilon_n \log(\pi(K_0)) + \liminf_{n\to\infty} \varepsilon_n \log\left(\frac{1}{\pi(K_0)}\int_{K_0} P_\gamma(X_{\varepsilon_n}\in C)\pi(\mathrm{d}\gamma)\right)$$

$$\geqslant \liminf_{n\to\infty} \frac{1}{\pi(K_0)}\int_{K_0} \varepsilon_n \log P_\gamma(X_{\varepsilon_n}\in C)\pi(\mathrm{d}\gamma).$$

By (6.8),

$$\geqslant \frac{1}{\pi(K_0)}\int_{K_0} (-(\lambda+\delta)^2)\pi(\mathrm{d}\gamma) = -(\lambda+\delta)^2.$$

Since δ is arbitrary, the theorem is proved. □

Acknowledgements

I am grateful to M. Röckner, A. Schied and J. Ubøe for stimulating discussions during the preparation of this work. My special thanks go to L. Gross for the stimulating discussions and his hospitality during my stay at Cornell University. Actually, the proof of Proposition 4.4 is based on a discussion with him. I also thank the anonymous referee for the very careful reading and useful comments.

References

Albeverio, S., Kondratiev, Yu.G., Röckner, M., 1996a. Differential geometry of poisson spaces. C.R. Acad. Sci. Paris 323, 1129–1134.

Albeverio, S., Kondratiev, Yu.G., Röckner, M., 1996b. Canonical Dirichlet operator and distorted Brownian motion on Poisson spaces. C.R. Acad. Sci. Paris 323, 1179–1184.

Albeverio, S., Kondratiev, Yu.G., Röckner, M., 1997. Analysis and geometry on configuration spaces. Bielefeld, preprint.

Aida, S., Zhang, T.S., 1999. On the small time asymptotics of diffusion processes on path groups, preprint. Potential Anal., in preparation.

Dembo, A., Zeitouni, O., 1992. Large Deviations Techniques and Applications. Jones and Bartlett Publishers, Boston.

Fukushima, M., Oshima, Y., Takeda, M., 1994. Dirichlet Forms and Symmetric Markov Processes. Walter de Gruyter, Berlin.

Fang, S., 1994. On the Ornstein–Uhlenbeck process. Stochastics Stochastic Rep. 46, 141–159.

Fang, S., Zhang, T.S., 1999. On the small time behavior of Ornstein–Uhlenbeck processes with unbounded linear drifts. Probab. Theory Related Fields 114 (4), 487–504.

Gross, L., 1980. Thermodynamics, Statistical Mechanics and Random Fields, Ecole d'Ete de Probabilites de Saint-Flour X-1980, Lecture Notes in Mathematics, Vol. 929. Springer, Berlin.

Lyons, T.J., Zhang, T.S., 1996. Convergence of non-symmetric Dirichlet processes. Stochastics Stochastic Rep. 57, 159–167.

Lyons, T., Zheng, W.A., 1988. Acrossing estimate for the canonical process on a Dirichlet space and tightness result. Colloque Paul Levy sur Les processes stochastique. Asterisque 157–158, 248–272.

Ma, Z.M., Röckner, M., 1992. Introduction to the Theory of (Non.Symmetric) Dirichlet Forms. Springer, Berlin.

Röckner, M., Schied, A., 1999. Rademacher's Theorem on configuration spaces and applications. J. Functional Anal. 169, 325–356.

Zhang, T.S., 2000. On the small time asymptotics of diffusion processes on Hilbert spaces, preprint. Ann. Probab. 28 (2).